

#22  
10-12-02  
5

1647



TECH CENTER 1600/2900

3 2002

RECEIVED

1600

## RAW SEQUENCE LISTING

DATE: 08/07/2002

PATENT APPLICATION: US/09/542,520B

TIME: 15:57:24

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

P.5

ENTERED

5 <110> APPLICANT: Jackson, W. James  
6 Pace, John  
8 <120> TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
10 <130> FILE REFERENCE: 7969-076-999  
12 <140> CURRENT APPLICATION NUMBER: 09/542,520B  
13 <141> CURRENT FILING DATE: 2000-04-03  
15 <150> PRIOR APPLICATION NUMBER: 08/942,596  
16 <151> PRIOR FILING DATE: 1997-10-02  
18 <160> NUMBER OF SEQ ID NOS: 43  
20 <170> SOFTWARE: PatentIn version 3.0  
22 <210> SEQ ID NO: 1  
23 <211> LENGTH: 4435  
24 <212> TYPE: DNA  
25 <213> ORGANISM: Artificial Sequence  
27 <220> FEATURE:  
28 <223> OTHER INFORMATION: Description of Artificial Sequence: Recombinant Expression  
Vector  
30 <220> FEATURE:  
31 <221> NAME/KEY: CDS  
32 <222> LOCATION: (382)..(3417)  
33 <223> OTHER INFORMATION:  
35 <400> SEQUENCE: 1  
36 gggcaaaact cttccccccg ggatttatat gggaaagggg aaacttttggc ccgtattcaa 60  
37 gcgccacggg ttttggggcg gaatgaattt tttcgttccg gaaaaagtaa ttccccggga 120  
38 acgtagggta tcggtttcat aggcctcgcca aatgggatat aggtggaaag gtaaaaaaaa 180  
39 ctgagccaag caaaggatag agaagtcttg taatcatcgc aggttaaagg ggggatggtta 240  
40 ttttagcctg caaatagtgt aattattgga tcctgttaaag agaaaaggac gaatgcgctg 300  
41 aagataagaa cattttattga tattaataa ttaatttttt atgaagcgga gtaattaatt 360  
42 ttatctctca gcttttgtgt g atg caa acg tct ttc cat aag ttc ttt ctt 411  
43 Met Gln Thr Ser Phe His Lys Phe Phe Leu  
44 1 5 10  
45 tca atg att cta gct tat tct tgc tgc tct tta aat ggg ggg gga tat 459  
47 Ser Met Ile Leu Ala Tyr Ser Cys Cys Ser Leu Asn Gly Gly Gly Tyr  
48 15 20 25  
49 gca gca gaa atc atg gtt cct caa gga att tac gat ggg gag acg tta 507  
50 Ala Ala Glu Ile Met Val Pro Gln Gly Ile Tyr Asp Gly Glu Thr Leu  
51 30 35 40  
52 act gta tca ttt ccc tat act gtt ata gga gat ccg agt ggg act act 555  
53 Thr Val Ser Phe Pro Tyr Thr Val Ile Gly Asp Pro Ser Gly Thr Thr  
54 45 50 55  
55 gtt ttt tct gca gga gag tta aca tta aaa aat ctt gac aat tct att 603  
56 Val Phe Ser Ala Gly Glu Leu Thr Leu Lys Asn Leu Asp Asn Ser Ile  
57 60 65 70  
58 gca gct ttg cct tta agt tgt ttt ggg aac tta tta ggg agt ttt act 651

## RAW SEQUENCE LISTING

DATE: 08/07/2002

PATENT APPLICATION: US/09/542,520B

TIME: 15:57:24

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

59	Ala	Ala	Leu	Pro	Leu	Ser	Cys	Phe	Gly	Asn	Leu	Leu	Gly	Ser	Phe	Thr	
60	75					80				85						90	
61	ggt	tta	ggg	aga	gga	cac	tcg	ttg	act	ttc	gag	aac	ata	cgg	act	tct	699
62	Val	Leu	Gly	Arg	Gly	His	Ser	Leu	Thr	Phe	Glu	Asn	Ile	Arg	Thr	Ser	
63					95					100					105		
64	aca	aat	ggg	gca	gct	cta	agt	aat	agc	gct	gct	gat	gga	ctg	ttt	act	747
65	Thr	Asn	Gly	Ala	Ala	Leu	Ser	Asn	Ser	Ala	Ala	Asp	Gly	Leu	Phe	Thr	
66				110						115				120			
67	att	gag	ggt	ttt	aaa	gaa	tta	tcc	ttt	tcc	aat	tgc	aat	tca	tta	ctt	795
68	Ile	Glu	Gly	Phe	Lys	Glu	Leu	Ser	Phe	Ser	Asn	Cys	Asn	Ser	Leu	Leu	
69		125						130					135				
70	gcc	gta	ctg	cct	gct	gca	acg	act	aat	aag	ggt	agc	cag	act	ccg	acg	843
71	Ala	Val	Leu	Pro	Ala	Ala	Thr	Thr	Asn	Lys	Gly	Ser	Gln	Thr	Pro	Thr	
72		140					145						150				
73	aca	aca	tct	aca	ccg	tct	aat	ggt	act	att	tat	tct	aaa	aca	gat	ctt	891
74	Thr	Thr	Ser	Thr	Pro	Ser	Asn	Gly	Thr	Ile	Tyr	Ser	Lys	Thr	Asp	Leu	
75	155					160				165					170		
76	ttg	tta	ctc	aat	aat	gag	aag	ttc	tca	ttc	tat	agt	aat	tta	gtc	tct	939
77	Leu	Leu	Leu	Asn	Asn	Glu	Lys	Phe	Ser	Phe	Tyr	Ser	Asn	Leu	Val	Ser	
78				175						180				185			
79	gga	gat	ggg	gga	gct	ata	gat	gct	aag	agc	tta	acg	ggt	caa	gga	att	987
80	Gly	Asp	Gly	Gly	Ala	Ile	Asp	Ala	Lys	Ser	Leu	Thr	Val	Gln	Gly	Ile	
81			190						195					200			
82	agc	aag	ctt	tgt	gtc	ttc	caa	gaa	aat	act	gct	caa	gct	gat	ggg	gga	1035
83	Ser	Lys	Leu	Cys	Val	Phe	Gln	Glu	Asn	Thr	Ala	Gln	Ala	Asp	Gly	Gly	
84			205					210					215				
85	gct	tgt	caa	gta	gtc	acc	agt	ttc	tct	gct	atg	gct	aac	gag	gct	cct	1083
86	Ala	Cys	Gln	Val	Val	Thr	Ser	Phe	Ser	Ala	Met	Ala	Asn	Glu	Ala	Pro	
87		220					225						230				
88	att	gcc	ttt	gta	gcg	aat	gtt	gca	gga	gta	aga	ggg	gga	ggg	att	gct	1131
89	Ile	Ala	Phe	Val	Ala	Asn	Val	Ala	Gly	Val	Arg	Gly	Gly	Gly	Ile	Ala	
90	235					240				245					250		
91	gct	gtt	cag	gat	ggg	cag	cag	gga	gtg	tca	tca	tct	act	tca	aca	gaa	1179
93	Ala	Val	Gln	Asp	Gly	Gln	Gln	Gly	Val	Ser	Ser	Ser	Thr	Ser	Thr	Glu	
94				255						260				265			
95	gat	cca	gta	gta	agt	ttt	tcc	aga	aat	act	gcg	gta	gag	ttt	gat	ggg	1227
96	Asp	Pro	Val	Val	Ser	Phe	Ser	Arg	Asn	Thr	Ala	Val	Glu	Phe	Asp	Gly	
97			270						275					280			
98	aac	gta	gcc	cga	gta	gga	gga	ggg	att	tac	tcc	tac	ggg	aac	gtt	gct	1275
99	Asn	Val	Ala	Arg	Val	Gly	Gly	Gly	Ile	Tyr	Ser	Tyr	Gly	Asn	Val	Ala	
100		285						290					295				
101	ttc	ctg	aat	aat	gga	aaa	acc	ttg	ttt	ctc	aac	aat	gtt	gct	tct	cct	1323
102	Phe	Leu	Asn	Asn	Gly	Lys	Thr	Leu	Phe	Leu	Asn	Asn	Val	Ala	Ser	Pro	
103		300					305						310				
104	gtt	tac	att	gct	gct	aag	caa	cca	aca	agt	gga	cag	gct	tct	aat	acg	1371
105	Val	Tyr	Ile	Ala	Ala	Lys	Gln	Pro	Thr	Ser	Gly	Gln	Ala	Ser	Asn	Thr	
106	315					320					325				330		
107	agt	aat	aat	tac	gga	gat	gga	gga	gct	atc	ttc	tgt	aag	aat	ggt	gcg	1419
108	Ser	Asn	Asn	Tyr	Gly	Asp	Gly	Gly	Ala	Ile	Phe	Cys	Lys	Asn	Gly	Ala	

## RAW SEQUENCE LISTING

DATE: 08/07/2002

PATENT APPLICATION: US/09/542,520B

TIME: 15:57:24

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

109		335		340		345		
110	caa gca gga tcc aat aac tct gga tca gtt tcc ttt gat gga gag gga							1467
111	Gln Ala Gly Ser Asn Asn Ser Gly Ser Val Ser Phe Asp Gly Glu Gly							
112		350		355		360		
113	gta gtt ttc ttt agt agc aat gta gct gct ggg aaa ggg gga gct att							1515
114	Val Val Phe Phe Ser Ser Asn Val Ala Ala Gly Lys Gly Gly Ala Ile							
115		365		370		375		
116	tat gcc aaa aag ctc tct gtt gct aac tgt ggc cct gta caa ttt tta							1563
117	Tyr Ala Lys Lys Leu Ser Val Ala Asn Cys Gly Pro Val Gln Phe Leu							
118		380		385		390		
119	agg aat atc gct aat gat ggt gga gcg att tat tta gga gaa tct gga							1611
120	Arg Asn Ile Ala Asn Asp Gly Gly Ala Ile Tyr Leu Gly Glu Ser Gly							
121	395		400		405		410	
122	gag ctc agt tta tct gct gat tat gga gat att att ttc gat ggg aat							1659
123	Glu Leu Ser Leu Ser Ala Asp Tyr Gly Asp Ile Ile Phe Asp Gly Asn							
124		415		420		425		
130	ctt aaa aga aca gcc aaa gag aat gct gcc gat gtt aat ggc gta act							1707
131	Leu Lys Arg Thr Ala Lys Glu Asn Ala Ala Asp Val Asn Gly Val Thr							
132		430		435		440		
133	gtg tcc tca caa gcc att tct atg gga tct gga ggg aaa ata acg aca							1755
134	Val Ser Ser Gln Ala Ile Ser Met Gly Ser Gly Gly Lys Ile Thr Thr							
135		445		450		455		
136	tta aga gct aaa gca ggg cat cag att ctc ttt aat gat ccc atc gag							1803
137	Leu Arg Ala Lys Ala Gly His Gln Ile Leu Phe Asn Asp Pro Ile Glu							
139		460		465		470		
140	atg gca aac gga aat aac cag cca gcg cag tct tcc aaa ctt cta aaa							1851
141	Met Ala Asn Gly Asn Asn Gln Pro Ala Gln Ser Ser Lys Leu Leu Lys							
142	475		480		485		490	
143	att aac gat ggt gaa gga tac aca ggg gat att gtt ttt gct aat gga							1899
144	Ile Asn Asp Gly Glu Gly Tyr Thr Gly Asp Ile Val Phe Ala Asn Gly							
145		495		500		505		
146	agc agt act ttg tac caa aat gtt acg ata gag caa gga agg att gtt							1947
147	Ser Ser Thr Leu Tyr Gln Asn Val Thr Ile Glu Gln Gly Arg Ile Val							
148		510		515		520		
149	ctt cgt gaa aag gca aaa tta tca gtg aat tct cta agt cag aca ggt							1995
150	Leu Arg Glu Lys Ala Lys Leu Ser Val Asn Ser Leu Ser Gln Thr Gly							
151		525		530		535		
152	ggg agt ctg tat atg gaa gct ggg agt aca tgg gat ttt gta act cca							2043
153	Gly Ser Leu Tyr Met Glu Ala Gly Ser Thr Trp Asp Phe Val Thr Pro							
154		540		545		550		
155	caa cca cca caa cag cct cct gcc gct aat cag ttg atc acg ctt tcc							2091
156	Gln Pro Pro Gln Gln Pro Pro Ala Ala Asn Gln Leu Ile Thr Leu Ser							
157	555		560		565		570	
158	aat ctg cat ttg tct ctt tct tct ttg tta gca aac aat gca gtt acg							2139
159	Asn Leu His Leu Ser Leu Ser Ser Leu Leu Ala Asn Asn Ala Val Thr							
160		575		580		585		
161	aat cct cct acc aat cct cca gcg caa gat tct cat cct gca gtc att							2187
162	Asn Pro Pro Thr Asn Pro Pro Ala Gln Asp Ser His Pro Ala Val Ile							
163		590		595		600		

## RAW SEQUENCE LISTING

DATE: 08/07/2002

PATENT APPLICATION: US/09/542,520B

TIME: 15:57:24

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

164	ggt	agc	aca	act	gct	ggt	tct	ggt	aca	att	agt	ggg	cct	atc	ttt	ttt	2235
165	Gly	Ser	Thr	Thr	Ala	Gly	Ser	Val	Thr	Ile	Ser	Gly	Pro	Ile	Phe	Phe	
166			605					610					615				
167	gag	gat	ttg	gat	gat	aca	gct	tat	gat	agg	tat	gat	tgg	cta	ggt	tct	2283
168	Glu	Asp	Leu	Asp	Asp	Thr	Ala	Tyr	Asp	Arg	Tyr	Asp	Trp	Leu	Gly	Ser	
169		620					625					630					
170	aat	caa	aaa	atc	aat	gtc	ctg	aaa	tta	cag	tta	ggg	act	aag	ccc	cca	2331
171	Asn	Gln	Lys	Ile	Asn	Val	Leu	Lys	Leu	Gln	Leu	Gly	Thr	Lys	Pro	Pro	
172	635					640					645					650	
173	gct	aat	gcc	cca	tca	gat	ttg	act	cta	ggg	aat	gag	atg	cct	aag	tat	2379
174	Ala	Asn	Ala	Pro	Ser	Asp	Leu	Thr	Leu	Gly	Asn	Glu	Met	Pro	Lys	Tyr	
175				655						660					665		
176	ggc	tat	caa	gga	agc	tgg	aag	ctt	gcg	tgg	gat	cct	aat	aca	gca	aat	2427
177	Gly	Tyr	Gln	Gly	Ser	Trp	Lys	Leu	Ala	Trp	Asp	Pro	Asn	Thr	Ala	Asn	
178			670					675						680			
179	aat	ggt	cct	tat	act	ctg	aaa	gct	aca	tgg	act	aaa	act	ggg	tat	aat	2475
180	Asn	Gly	Pro	Tyr	Thr	Leu	Lys	Ala	Thr	Trp	Thr	Lys	Thr	Gly	Tyr	Asn	
181			685				690						695				
182	cct	ggg	cct	gag	cga	gta	gct	tct	ttg	ggt	cca	aat	agt	tta	tgg	gga	2523
183	Pro	Gly	Pro	Glu	Arg	Val	Ala	Ser	Leu	Val	Pro	Asn	Ser	Leu	Trp	Gly	
185		700				705					710						
186	tcc	att	tta	gat	ata	cga	tct	gcg	cat	tca	gca	att	caa	gca	agt	gtg	2571
187	Ser	Ile	Leu	Asp	Ile	Arg	Ser	Ala	His	Ser	Ala	Ile	Gln	Ala	Ser	Val	
188	715				720					725					730		
189	gat	ggg	cgc	tct	tat	tgt	cga	gga	tta	tgg	ggt	tct	gga	ggt	tcg	aat	2619
190	Asp	Gly	Arg	Ser	Tyr	Cys	Arg	Gly	Leu	Trp	Val	Ser	Gly	Val	Ser	Asn	
191				735						740					745		
192	ttc	ttc	tat	cat	gac	cgc	gat	gct	tta	ggt	cag	gga	tat	cgg	tat	att	2667
193	Phe	Phe	Tyr	His	Asp	Arg	Asp	Ala	Leu	Gly	Gln	Gly	Tyr	Arg	Tyr	Ile	
194			750					755						760			
195	agt	ggg	ggt	tat	tcc	tta	gga	gca	aac	tcc	tac	ttt	gga	tca	tcg	atg	2715
196	Ser	Gly	Gly	Tyr	Ser	Leu	Gly	Ala	Asn	Ser	Tyr	Phe	Gly	Ser	Ser	Met	
197			765				770						775				
198	ttt	ggt	cta	gca	ttt	acc	gaa	gta	ttt	ggt	aga	tct	aaa	gat	tat	gta	2763
199	Phe	Gly	Leu	Ala	Phe	Thr	Glu	Val	Phe	Gly	Arg	Ser	Lys	Asp	Tyr	Val	
200		780				785						790					
201	gtg	tgt	cgt	tcc	aat	cat	cat	gct	tgc	ata	gga	tcc	ggt	tat	cta	tct	2811
202	Val	Cys	Arg	Ser	Asn	His	His	Ala	Cys	Ile	Gly	Ser	Val	Tyr	Leu	Ser	
203	795				800						805				810		
204	acc	caa	caa	gct	tta	tgt	gga	tcc	tat	ttg	ttc	gga	gat	gcg	ttt	atc	2859
205	Thr	Gln	Gln	Ala	Leu	Cys	Gly	Ser	Tyr	Leu	Phe	Gly	Asp	Ala	Phe	Ile	
206				815						820					825		
207	cgt	gct	agc	tac	ggg	ttt	ggg	aat	cag	cat	atg	aaa	acc	tca	tat	aca	2907
208	Arg	Ala	Ser	Tyr	Gly	Phe	Gly	Asn	Gln	His	Met	Lys	Thr	Ser	Tyr	Thr	
209				830						835					840		
210	ttt	gca	gag	gag	agc	gat	ggt	cgt	tgg	gat	aat	aac	tgt	ctg	gct	gga	2955
211	Phe	Ala	Glu	Glu	Ser	Asp	Val	Arg	Trp	Asp	Asn	Asn	Cys	Leu	Ala	Gly	
212			845				850						855				
213	gag	att	gga	gcg	gga	tta	ccg	att	gtg	att	act	cca	tct	aag	ctc	tat	3003

## RAW SEQUENCE LISTING

DATE: 08/07/2002

PATENT APPLICATION: US/09/542,520B

TIME: 15:57:24

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

```

214 Glu Ile Gly Ala Gly Leu Pro Ile Val Ile Thr Pro Ser Lys Leu Tyr
215      860      865      870
216 ttg aat gag ttg cgt cct ttc gtg caa gct gag ttt tct tat gcc gat      3051
217 Leu Asn Glu Leu Arg Pro Phe Val Gln Ala Glu Phe Ser Tyr Ala Asp
218 875      880      885      890
219 cat gaa tct ttt aca gag gaa ggc gat caa gct cgg gca ttc aag agc      3099
220 His Glu Ser Phe Thr Glu Glu Gly Asp Gln Ala Arg Ala Phe Lys Ser
221      895      900      905
222 gga cat ctc cta aat cta tca gtt cct gtt gga gtg aag ttt gat cga      3147
223 Gly His Leu Leu Asn Leu Ser Val Pro Val Gly Val Lys Phe Asp Arg
224      910      915      920
225 tgt tct agt aca cat cct aat aaa tat agc ttt atg gcg gct tat atc      3195
226 Cys Ser Ser Thr His Pro Asn Lys Tyr Ser Phe Met Ala Ala Tyr Ile
227      925      930      935
228 tgt gat gct tat cgc acc atc tct ggt act gag aca acg ctc cta tcc      3243
229 Cys Asp Ala Tyr Arg Thr Ile Ser Gly Thr Glu Thr Thr Leu Leu Ser
231      940      945      950
232 cat caa gag aca tgg aca aca gat gcc ttt cat tta gca aga cat gga      3291
233 His Gln Glu Thr Trp Thr Thr Asp Ala Phe His Leu Ala Arg His Gly
234 955      960      965      970
235 gtt gtg gtt aga gga tct atg tat gct tct cta aca agt aat ata gaa      3339
236 Val Val Val Arg Gly Ser Met Tyr Ala Ser Leu Thr Ser Asn Ile Glu
237      975      980      985
238 gta tat ggc cat gga aga tat gag tat cga gat gct tct cga ggc tat      3387
239 Val Tyr Gly His Gly Arg Tyr Glu Tyr Arg Asp Ala Ser Arg Gly Tyr
240      990      995      1000
241 ggt ttg agt gca gga agt aga gtc cgg ttc taaaaatatt ggtagatag      3437
242 Gly Leu Ser Ala Gly Ser Arg Val Arg Phe
243      1005      1010
244 ttaagtgtta gcgatgcott tttctttgag atctacatca ttttgttttt tagcttgttt      3497
245 gtgttccctat tcgtatggat tcgcgagctc tctcaagtg ttaacgccta atgtaaccac      3557
246 tctttttaag ggagacgatg tttacttgaa tggagactgc gcttttgatca atgtctatgc      3617
247 aggagctgaa gaaggttcga ttatctcagc taatggcgac aatttaacga ttaccggaca      3677
248 aaaccataca ttatcattta cagattctca agggccagtt cttcaaaatt atgccttcac      3737
249 ttcagcagga gagacactta ctctgagaga ttttctgagt ctgatgttct cgaaaaatgt      3797
250 ttcttgcgga gaaaagggaa tgatctccgg gaaaaccgtg agtatttccg gagcaggcga      3857
251 agtgattttc tgggataact ccgtggggtg ttctccttta tctactgtgc caacctcatc      3917
252 atcaactccg cctgctccaa cagttagtga tgctcggaag ggggtctattt tttctgtaga      3977
253 gactagtttg gagatctcag gcgtcaaaaa aggggtcatg ttcgataata atgccgggaa      4037
254 tttcggaaca gtttttcgag gtaagaataa taataatgct ggtggtggag gcagtgggtt      4097
257 ccgtacacc atcaagtacg acttttacag ttaaaaactg taaagggaaa gtttctttca      4157
258 cagataacgt agcctcttgc ggaggcggag tgggttataa aggcattgtg cttttcaaag      4217
259 acaatgaagg aggcattatc ttccgagga acacagcata cgatgattta aggattcttg      4277
260 ctgctactaa tcaggatcag aatacggaga caggaggcgg tggaggagtt atttgcctc      4337
261 cagatgattc tgtaaagttt gaaggcaata aaggttctat tgtttttgat tacaactttg      4397
262 caaaaggcag aggcggaagc atcctaacga aagaattc      4435
264 <210> SEQ ID NO: 2
265 <211> LENGTH: 1012
266 <212> TYPE: PRT

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
Review the Sequence Listing to insure a corresponding  
explanation is presented in the <220> to <223> fields of  
each sequence using n or Xaa.

psl

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/542,520B

DATE: 08/07/2002

TIME: 15:57:25

Input Set : A:\EP.txt

Output Set: N:\CRF3\08072002\I542520B.raw

L:415 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:429 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:442 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:455 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:469 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:482 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8  
L:495 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9